$$0 \leq Pr[a_k] \leq 1$$

(6.48)

$$\sum_{k=1}^{K} (Pr[a_k]) = 1$$

(6.49)

This coin-flipping model assumes that symbols occur without regard to what preceding or succeeding symbols were, a false assumption for typed text. Despite this probabilistic model's over-simplicity, the ideas we develop here also work when more accurate, but still probabilistic, models are used. The key quantity that characterizes a symbolic-valued signal is the **entropy** of its alphabet.

$$H(A) = -\left(\sum_{k} (Pr[a_k] \log_2 (Pr[a_k]))\right)$$

Because we use the base-2 logarithm, entropy has units of bits. For this Definition to make sense, we must take special note of symbols having probability zero of occurring. A zero-probability symbol never occurs; thus, we define $0\log_2 0=0$ so that such symbols do not affect the entropy. The maximum value attainable by an alphabet's entropy occurs when the symbols are equally likely

$$(Pr[a_k] = Pr[a_l]) .$$

In this case, the entropy equals $\log_2 K$. The minimum value occurs when only one symbol occurs; it has probability one of occurring and the rest have probability zero.

Exercise 6.20.1

Derive the maximum-entropy results, both the numeric aspect (entropy equals $\log_2 K$) and the theoretical one (equally likely symbols maximize entropy). Derive the value of the minimum entropy alphabet.

Example 6.1

A four-symbol alphabet has the following probabilities.

$$Pr[a_0] = \frac{1}{2}$$
$$Pr[a_1] = \frac{1}{4}$$
$$Pr[a_2] = \frac{1}{8}$$
$$Pr[a_3] = \frac{1}{8}$$

Note that these probabilities sum to one as they should. As

$$\frac{1}{2} = 2^{-1} , \log_2 \left(\frac{1}{2}\right) = -1.$$

The entropy of this alphabet equals